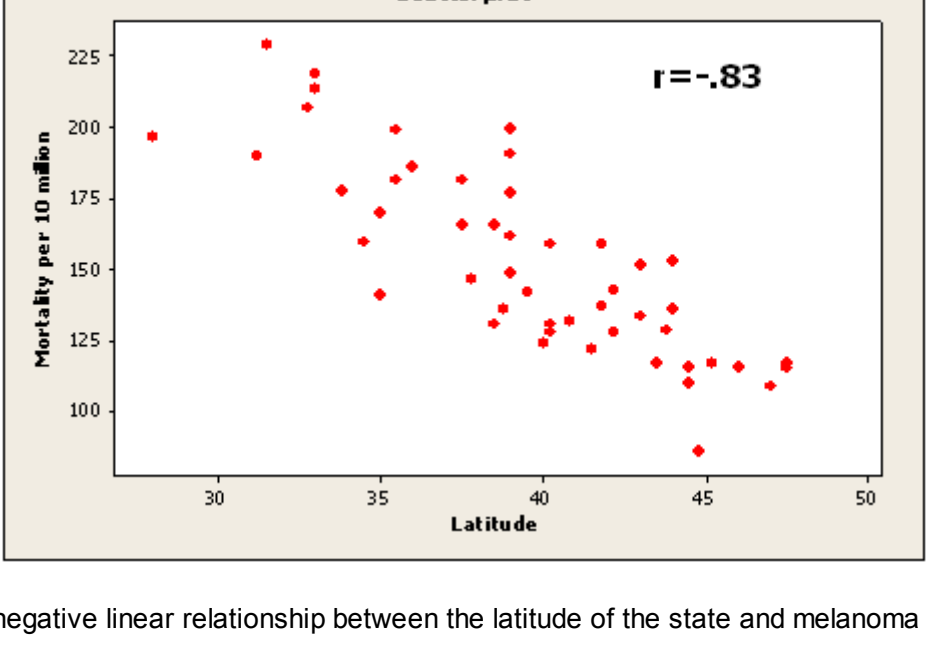


# Solutions: Case C→C and Q→Q Checkpoint

The first three questions refer to the following information:

The mortality rate from melanoma (skin cancer) during the 1950s was recorded for each of the 48 contiguous United States, plus Washington D.C. (as reported by Fisher and Van Belle (1993) and found on <http://www.stat.psu.edu/~lsimon/stat501wc/sp05/data/>).

The following is the scatterplot of the data:



The plot shows a negative linear relationship between the latitude of the state and melanoma mortality rate.

## Question 1

- In context, the negative relationship (shown in the scatterplot above) means that:
- ☐ (a) The more southern the state, the higher the melanoma death rate.
  - ☐ (b) The more northern the state, the higher the melanoma death rate.
  - ☐ (c) The more southern the state, the more people died of melanoma.

Correct answer: (a)

Select one answer.  
10 points

## Question 2

- Based on the scatterplot and the value of the correlation coefficient, it would make sense to test the significance of this observed linear relationship between latitude and melanoma mortality rate. The appropriate hypotheses are:
- ☐ (a)  
 $H_0$ : melanoma mortality rate is related to latitude  
 $H_a$ : melanoma mortality rate is not related to latitude
  - ☐ (b)  
 $H_0$ : melanoma mortality rate is not related to latitude  
 $H_a$ : melanoma mortality rate is related to latitude
  - ☐ (c)  
 $H_0$ : melanoma mortality rate is linearly related to latitude  
 $H_a$ : melanoma mortality rate is not linearly related to latitude
  - ☐ (d)  
 $H_0$ : melanoma mortality rate is not linearly related to latitude  
 $H_a$ : melanoma mortality rate is linearly related to latitude

Correct answer: (d)

Select one answer.  
10 points

## Question 3

- The following output is available:
- Regression Analysis: Mort versus Lat**
- The regression equation is  
Mort = 389 - 5.98 Lat
- | Predictor | Coef    | SE Coef | T     | P     |
|-----------|---------|---------|-------|-------|
| Constant  | 389.19  | 23.81   | 16.34 | 0.000 |
| Lat       | -5.9776 | 0.5984  | -9.99 | 0.000 |

- From the output we learn that:
- ☐ (a) the data do not provide sufficient evidence to conclude that melanoma mortality rate is linearly related to latitude.
  - ☐ (b) the data provide extremely strong evidence that melanoma mortality rate is linearly related to latitude.
  - ☐ (c) the data provide moderately strong evidence that melanoma mortality rate is linearly related to latitude.

Correct answer: (b)

Select one answer.  
10 points

The following four questions (i.e., **questions 18-21**) refer to the following information:

To determine if there is a relationship between grade performance and extracurricular participation, North Carolina state conducted a study of 112 students, recording the number of students in each of three extra-curricular categories, and each of two grade categories. Here are the results:

	Good grades	Poor grades
Low extra-curricular participation	11	2
Moderate extra-curricular participation	68	23
High extra-curricular participation	3	5

(Source: Felder, et. al., "A Study of Student Performance in an Introductory Chemical Engineering Course," 1992 ASEE Annual Conference Proceedings, pp. 1516–1519.)

## Question 4

- What is the expected count of students with high extra-curricular participation and good grades?
- ☐ (a) 3
  - ☐ (b) 5.86
  - ☐ (c) 8
  - ☐ (d) 112
  - ☐ (e) None of the above.

Correct answer: (b)

Select one answer.  
10 points

## Question 5

- The count that you found in the previous question is the number of students with high extra-curricular participation and good grades that you would expect to see assuming that:
- ☐ (a) the conditions that allow us to safely use the chi-squared procedure are met.
  - ☐ (b) extra-curricular participation and grade performance are related.
  - ☐ (c) extra-curricular participation and grade performance are independent.
  - ☐ (d) the null hypothesis of the chi-squared test for independence in this case is true.
  - ☐ (e) both (c) and (d) are correct.

Correct answer: (e)

Select one answer.  
10 points

## Question 6

- The following is the (edited) output of the chi-squared test in this case:
- Chi-Square Test: Good, Poor**
- Expected counts are printed below observed counts  
Chi-Square contributions are printed below expected counts
- |          | Good                 | Poor                 | Total |
|----------|----------------------|----------------------|-------|
| Low      | 11<br>9.52<br>0.231  | 2<br>3.48<br>0.631   | 13    |
| Moderate | 68<br>66.63<br>0.028 | 23<br>24.38<br>0.078 | 91    |
| High     | 3<br>1.394           | 5<br>2.14<br>3.810   | 8     |
| Total    | 82                   | 30                   | 112   |
- Chi-Sq = 6.171, DF = 2, P-Value = 0.046

- Based on the output (and using the traditional significance level of 5%) we can determine that:
- ☐ (a) the data provide sufficient evidence to conclude that extra-curricular participation and grade performance are independent.
  - ☐ (b) the data do not provide sufficient evidence to conclude that extra-curricular participation and grade performance are independent.
  - ☐ (c) the data provide sufficient evidence to conclude that extra-curricular participation and grade performance are related.
  - ☐ (d) the data do not provide sufficient evidence to conclude that extra-curricular participation and grade performance are related.

Correct answer: (c)

Select one answer.  
10 points

## Question 7

- Which of the following facts should make you the most worried about the reliability of the results of the test in this case?
- ☐ (a) The 112 students were all students in a chemical engineering course and not a random sample from the entire student body.
  - ☐ (b) Two of the six observed counts are less than 5.
  - ☐ (c) Not all of the cells' contributions to the chi-squared statistic are greater than 1.
  - ☐ (d) Two of the six expected counts are less than 5.

Correct answer: (d)

Select one answer.  
10 points